

Coursera Capstone

IBM Applied Data Science

Capstone

***opening a new bar in the city of London &
looking for a happy place to live .***

By: Balaram

may 2020

Business Problem

- 1)The objective of this capstone project is to analyse and select the best locations in the city of London to start a family or living or moving in from other places
-
- 2).The best place in the city of london to start a “Bar”. Using the techniques learnt throughout the course I aim to solve the problem.

Dat

a

- Data required

- ☐ List of neighbourhoods in London
- ☐ Latitude and longitude coordinates of the neighbourhoods
- ☐ Venue data, particularly data related to Bar.
- ☐ Data on happiness index.

- Sources of data

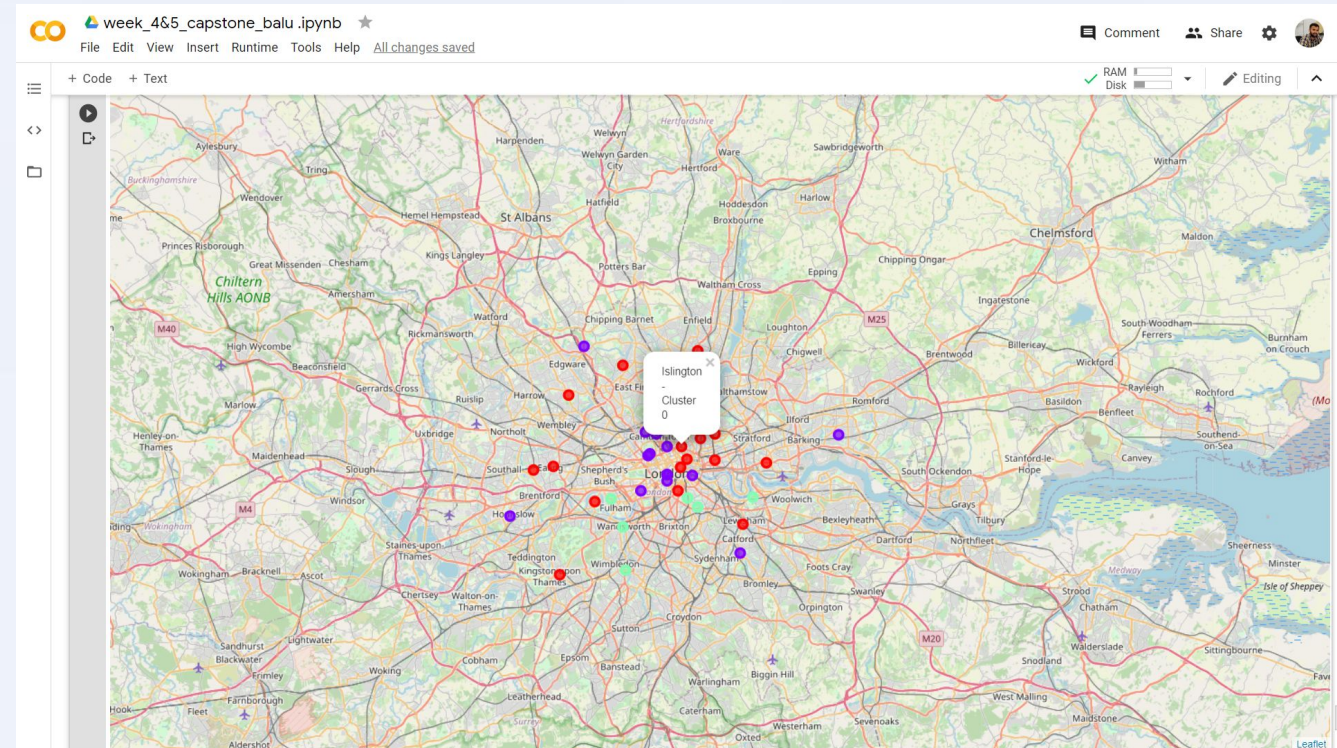
- ☐ london data store: burrough profiles.
- ☐ Geocoder package for latitude and longitude coordinates
- ☐ Foursquare API for venue data

Methodolog

- y Web scraping Wikipedia page for neighbourhoods list
 - Get latitude and longitude coordinates using Geocoder
 - map and cluster the happiness index data
 - Use Foursquare API to get venue data
 - Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
 - Filter venue category by Bar
 - Perform clustering on the data by using k-means clustering
 - Visualize the clusters in a map using Folium

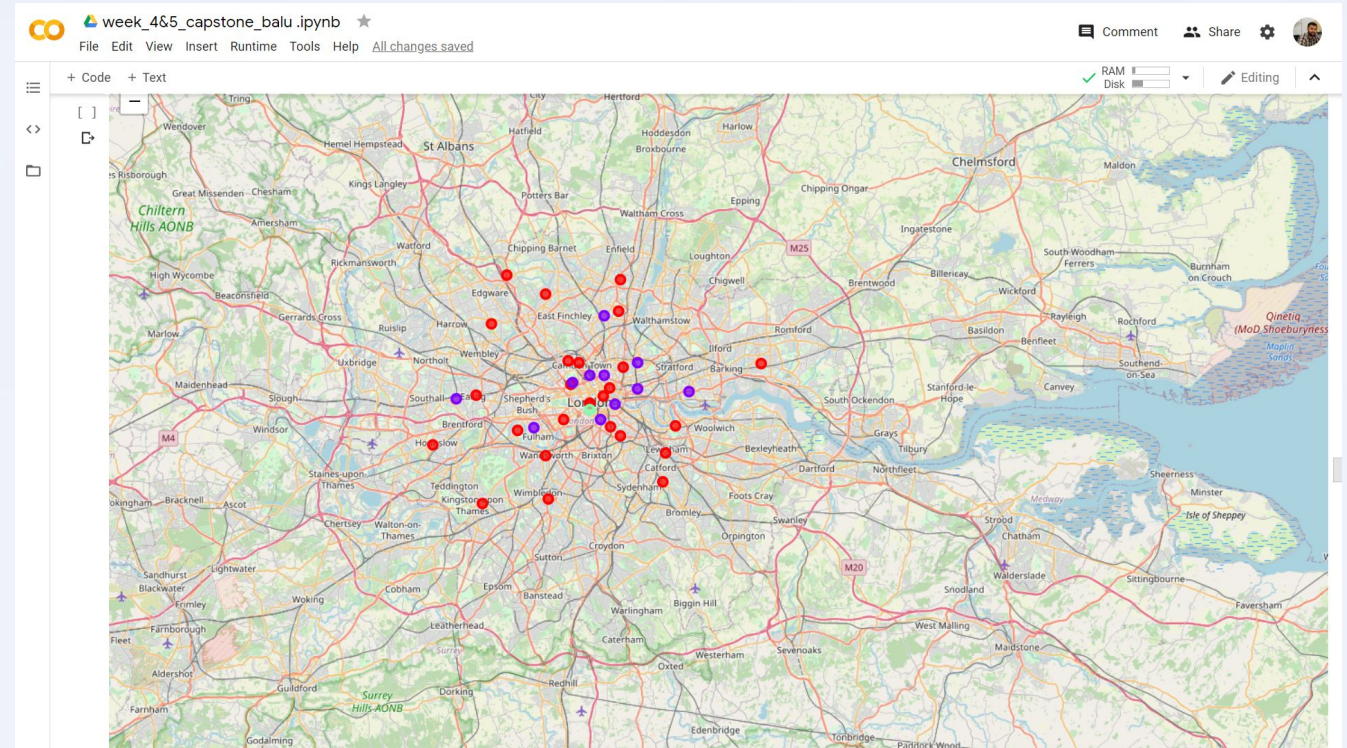
Result

- Categorized the neighbourhoods into 3 clusters :
 - ❑ Cluster 0: Neighbourhoods with high no of bars.
 - ❑ Cluster 1: Neighbourhoods with moderate number to no of bars
 - ❑ Cluster 2: Neighbourhoods with low concentration of shopping malls



Result

- S
- All the areas in the city of London ,England have been clustered into three regions as shown above in the map .They have been clustered based on the happiness index,worthwhileness index ,life satisfaction indices that give the overall living index of these areas.



Discussio

n

- The results from the k-means clustering of bars show that we can categorize the neighbourhoods into 3 clusters based on the frequency of occurrence for “bars”
- Cluster 0: Neighbourhoods with high number of bars
- Cluster 1: Neighbourhoods with moderate number to no existence of “bars”
- Cluster 2: Neighbourhoods with low concentration of bara

Recommendations

- Open new shopping malls in neighbourhoods in cluster 1 with little to no competition
- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of shopping malls and intense competition

Conclusio

n

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new shopping mall
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall

Thank you!

